DOCUMENT RESUME

ED 202 913 TM 810 396

AUTHOR Reed, Richard
TITLE The Results of a District-Wide Equating Study.
INSTITUTION Los Angeles Unified School District, Calif.
PUB DATE 14 Apr 81
NOTE 22p.; Paper presented at the Annual Meeting of the American Educational Research Association (65th, Los Angeles, CA, April 13-17, 1981).

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Elementary Education; *Equated Scores; Reading Achievement; *School Districts; *Testing Problems; *Testing Programs
IDENTIFIERS California Assessment Program; *Comprehensive Tests of Basic Skills; Equipercentile Equating; Los Angeles City Schools CA; *Survey of Essential Skills

ABSTRACT

The Los Angeles Unified School District has been concerned over the excess testing of elementary school pupils, particularly those who are in special programs. When the Board of Education decided to have yearly testing of pupil progress using a locally developed criterion-referenced test, the intention was to replace a previously used norm-referenced test. During the 1978-79 school year, a three-year equating study was begun. When it is completed, the District will be able to use its locally developed Survey of Essential Skills (SES) test in place of the Comprehensive Tests of Basic Skills (CTBS) testing for reporting Title I pupil progress. During the three-year period, all Title I pupils in grades 1 to 6 will be given both the CTBS and SES tests. The scores will be equated each year using the equipercentile method. The first two years of the study have been completed. This report examines the steps used and problems encountered. Preliminary findings suggest that all minimum conditions for a successful equating study have been met. (Author/GK)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED202913

LOS ANGELES UNIFIED SCHOOL DISTRICT
Research and Evaluation Branch

THE RESULTS OF A DISTRICT-WIDE EQUATING STUDY

Los Angeles, California
April 14, 1981

Prepared by:

Richard Reed

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

R. Reed

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☐ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

TM 810 396

THE SCOPE OF THIS PAPER

The concern of this paper is with testing at the elementary school level (grades 1-6). While there is also some concern with testing in the secondary schools, though less than in elementary, no plans have been made to change the present situation. Neither will this paper address the special testing of NES/LES nor special education pupils. When expressions such as "each pupil" or "all pupils" are used, they are referring to elementary school pupils who normally would be eligible for testing.

BACKGROUND

For several years, the Los Angeles Unified School district has been encouraged strongly to reduce the amount of mandated testing. This pressure has come from sources both within and without the District. Within the District the pressure has come mainly from local school personnel- teachers, principals, and the teachers' organizations. Outside of the District, a major sources of this pressure has been the State legislature. Senate Bill 1459, passed in 1980, set up a study to investigate the feasability of limiting State required testing to one test per pupil, per subject area, per year. Also, the Senate Committee on Education's hearings on teacher paperwork included a section on redundant local testing.

The present District mandated testing program consists of three tests given at the following grades in elementary school. The Comprehensive Tests of Basic Skills, Form S (CTBS), is given to all pupils in grades 3 and 5 and to Title I pupils in all grades. The Survey of Essential Skills (SES), a locally developed competency test, is given to all pupils in grades 1 to 6. The California Assessment Program (CAP), a matrix-sampling test which gives school and district information, is given in grades 1, 3, and 6. Table 1

gives the sections of the tests used and the amount of time required by each.

All of these tests to some extent reflect requirements to the district from State or Federal mandates. The CAP is a State testing program required of all schools. The CTBS is given for district and Title I reporting purposes. The SES is used to meet the State requirement for a competency test to be given at least once during grades 4 to 6.

There are three major complaints against "excessive" testing. One is that too much time is taken from the instructional program for testing and the necessary preparation. Secondly, there is the feeling that much of the testing is redundant, as all three tests measure achievement in reading and mathematics. Thirdly, the teachers feel there is little information of practical value on an individual pupil in either the CAP or CTBS; however, the teachers see some potential for program decisions from the newly developed SES.

With the implementation of model A2, which allows the use of non-standardized tests for reporting Title I progress, the District decided to use the SES to replace the CTBS for Title I reporting at grades 1 through 6.

OVERVIEW OF THE METHODOLOGY

In the spring of 1979, the first year SES was given, all pupils took the test, and a random sample of 2% of those scores was taken to be used for the equating study. The District followed its normal testing procedure for CTBS testing in that all pupils in grades 3 and 5 took the CTBS, while in grades 1, 2, and 4 only Title I pupils took it. In grade 6 all pupils took the reading part, but only Title I pupils took the mathematics.

This 2% random sample allowed for approximately 500 matched SES and

CTBS scores at grades 3 and 5 and at grade 6 reading. There were about 300 matched scores for grades 1, 2, and 4 and at grade 6 mathematics. This number was somewhat smaller than desired, but this particular year was a 'dry run' in that the SES was revised the following year and the original version will not be used. Equating was done by the equipercentile method and followed the ten steps listed below. Thes are, basically, the steps recommended by William Angoff in Thorndike's Educational Measurement, Second Edition.

1. The reading and mathematics scores from the SES tape were hand-matched with CTBS reading and mathematics scores for the same students. This allowed for 800 to 1000 matched scores for reading and mathematics at each grade level (for 1980 data).

2. The matched scores were run through SPSS subprograms FREQUENCIES and PEARSON CORR. This gave the cumulative percentiles necessary for equating and the correlation between tests. If all the statistics offered in the FREQUENCIES subprogram were requested, then much of the data necessary for calculating the standard error of equating were obtained also.

3. Relative cumulative frequencies (percentage of cases falling below each interval) were taken from the FREQUENCIES print-out for each raw score from zero to the highest possible score on a test. From these relative cumulative frequencies, mid-percentile ranks were calculated for both the SES and CTBS scores of the matched pairs.

4. Each percentile distribution was smoothed using the rolling weig ed average. If P(n) was the percentile of a score 'n' then P'(n), the smoothed percentile, was equal to $-3/35P(n-2)+12/35P(n-1)+17/35P(n)+12/35P(n+1)+(-3/35)P(n+2)$.

5. Each raw score on the SES was paired with a best estimate, to the nearest hundredth, CTBS score by matching mid-percentile ranks. Where exact matchings were not available, linear interpolation was used.

6. The paired score distribution was then smoothed, using the procedure from step 4.

7. The smoothed equated scores were rounded to the nearest whole number.

8. The internal reliability for each test was found. For the CTBS this was taken from the technical manual. For the SES this was done by using the SPSS subprogram RELIABILITY, model ALPHA, on the raw data supplied in the sample. This gave the Kuder-Richardson 20 reliability for the SES test.

9. The reliabilities for the paired tests were compared. They were considered equal if the square of the reliabilities (the amount of explained variance) differed by no more than .10. This was the case for all pairings.

10. The standard error of equating was calculated by the following formula, supplied by the Technical Assistance Center of the Northwest Regional Educational laboratory:

$$SE_y^2 = s_y^2 (1-r_{xy}) \frac{z_x^2 (1+r_{xy})+2}{N_t}$$

For the 1979-80 school year the random sample was increased to 5%. This allowed for 800 to 1000 matches at each grade level. The equating was done in the same manner as the previous year; however, the SES had been modified, so no comparison between years was made. The same modified version of the SES will be given during the 1980-81 school year, and the equating will be done again as a validation of the previous year. The District anticipates that this will be the last year in which all Title I pupils will take the CTBS.

Tables 2 through 7 at the back of this paper show the results of some of the steps in the equating. Grade 5 reading was chosen as a sample because it was representative of the typical equating results. These tables show the frequency distributions, equating results, national norm pairings, error of equating, and some necessary statistics to perform these operations.

SPECIFIC METHODOLOGICAL CONSIDERATIONS

The previous section gave a broad overview of the methodology of the equipercentile equating. This section will discuss five specific

considerations that might be of interest to any district which wished to become involved in equating. These five areas are the matching of scores, the calculation of equated scores, the effects of smoothing, the counter-balancing of testing, and the problems caused by the difference between the two tests.

Matching of Scores

In theory, the equipercentile equating of test scores does not require a sample of students who have taken both tests, only that both samples are representative of the same population. The matching of scores is necessary, however, for finding the correlation between the two tests. In this equating, only matched scores were used.

No student identification number system was available, and all past efforts at computer matching by name have proven unsuccessful; therefore, the matching was done by hand. Computer printouts were made with student SES scores for reading and mathematics by school and grade. These scores were hand matched with CTBS reading and mathematics scores from district rosters, which also were by grade and school.

This was the most time consuming step in the whole equating process. It is estimated that it took about 8 hours to hand match the 800 to 1000 pupils scores for each grade. On the other hand, once the computer programs had been developed, the actual equating for each test and grade could be done in about an hour.

Calculation of Equated Scores

Angoff suggests that not every score of the first test needs to be equated with a score on the second, but that certain representative points be equated and then the curve be drawn and smoothed to find the missing

paired scores. Since the district had access to the UNIVAC computer and FORTRAN programming, all scores on the SES were matched with a score on the CTBS. A district without a computer might not want to equate every score.

Smoothing

In the Angoff article, a smoothing of the cumulative distribution curves is recommended to eliminate any extraneous error that might occur. However, such smoothing is difficult and time-consuming if no computer is available; therefore, the personnel at the Technical Assistance Center of the Nortwest Regional Educational Laboratory (NWREL) do not include this step in their workshops on equating.

At the suggestion of the staff at NWREL, the district tried two equatings of the same tests, one using the smoothing and the other eliminating it. For 643 grade 5 pupils in reading, there was an average difference of.08 between a smoothed CTBS equated score and a non-smoothed score. The largest difference was .44. By coincidence, none of the rounded equated scores differed between the smoothed and non-smoothed distributions.

For 190 grade 2 pupils in reading, the average difference was .13 between the smoothed and non-smoothed scores. The largest difference was .48 and the smallest .02. There were, however, a few places where the rounded equated scores were different between the smoothed and non-smoothed distributions.

The result of the investigation suggested that smoothing made little difference for groups of this size and that the effect of smoothing is less as the size of the group increases.

Counterbalancing of the Testing

The recommended procedure for equating two tests when the same pupils take both of the tests is to counterbalance the testing. For equating two tests, A and B, a counterbalanced design calls for one group of students to take test A first and test B second. The other group of students take B first and A second. This design compensates for any possible effects of testing, where things learned on the first test help increase a pupil's score on the second.

The District did not counterbalance its equating. The testing of over 200,000 pupils is a complicated process in itself and requires testing schedules that are consistent across schools and set up well in advance of testing dates. To have reversed the testing order for any of the schools would have added difficulties to the testing program that would have more than outweighed the advantages of the counterbalanced design.

All pupils took the CTBS before the SES for the 1979-80 school year, leaving the possibility of a slightly inflated SES score from the testing effect. For the 1980-81 school year the order will be reversed, with the SES being given before the CTBS. This reversing of order over the two years will allow some investigation of any testing effects, but the findings will be confounded by some probable instruction effects as the teachers teach the skill continuums that are tested on the SES.

This testing effect should only be a problem in the first year of reporting for Title I. The spring 1980 test (with a slightly inflated score) will be used as a pretest, and the spring 1981 test (without the testing effect) will be used for the posttest. In subsequent years neither the pretest nor posttest will have the testing effect.

Differences Between the Two Tests

The CTBS is a nationally developed norm-referenced test with a normal curve distribution. The SES is a locally developed competency-based test. Its purpose is to measure the extent to which a pupil has learned basic skills from the District continuums. Ideally, a 100% score on the SES means only that a pupil has mastered what is expected of him, or is on grade level. The distribution of this test in negatively skewed.

Although the equipercentile method of equating allows for distributions of different shapes, not all the problems of these differing shapes are solved. The biggest problem in the equating was that the paired CTBS scores were not evenly distributed across the SES scores. Tables 7 and 8 show paired scores for grade 5 reading and grade 1 mathematics respectively. These were shown because grade 5 reading is a good sample of a typical equating, while grade 1 mathematics is an example of the worst situation.

Table 7 shows that a range of 43 CTBS scores, approximately half of the total possible, is spread across the first 35 SES scores (excluding scores of 0). This is somewhat equal; however, 42 CTBS scores (the total possible score is 85) are spread across the last 9 SES scores. Looking at it another way, there are 37 SES scores below the CTBS 50% national percentile, while there are only 7 scores above it.

The situation is worse at grade 1 mathematics, as table 8 shows. The district continuums are such that nearly all grade 1 pupils are able to meet the mathematics objectives for that grade. Half of the pupils miss no more than one question. As a result, there are 26 CTBS scores spread out across the first 26 SES scores, but 30 CTBS scores (for a total possible score of 56) spread out across that last 2 SES scores. Only the last two SES scores are above the 50th percentile nationally, and no pupil could get a national

percentile ranking above 81 on the SES even though he/she had a perfect score on the test.

This problem is at its worst for individual scores and for pupils whose achievement is above the 50th percentile nationally. As a result, individual pupil national percentile rankings will not be utilized by the District. Since Title I pupils have a median and mean which is below the 50th percentile nationally, this mean and median is not subject to as much of a problem.

The results of the equating study will be used only for reporting school and District Title I progress. No individual equated scores or percentiles will be reported. The primary purpose of the SES will remain the same- giving the teachers a diagnostic tool as to how well a pupil is mastering the district continuums. Pupil scores are reported to teachers and pupils on a percent correct basis.

FUTURE CONSIDERATIONS

The following are future problem areas identified by the equating study. In some cases decisions already have been made about handling the problem. In other cases the decisions have not yet been made.

New Form of the SES

The District anticipates having a second form of the SES ready by the 1981-82 school year. It, basically, will be an equivalent form to the present one; however, there may be some improvements, particularly in those tests in which a high percentage of pupils were getting nearly all items correct. The present plan is to equate that new form with the new CTBS, Form U, by using a sample of 800 to 1000 pupils at each grade who take both tests. The CTBS, Form S, equating to the CTBS, Form U, which will be done

by the publisher, will allow the District to equate scores from both forms of the SES with either form of the CTBS.

Future Up-dates on the Equating

Although, theoretically, the equating should not have to be repeated as long as the two tests remain the same, some up-dating may be necessary in this case. If the SES is responsive to the effects of the teaching of the District continuums, as is hoped, and the CTBS is not responsive, as expected, then one might guess that SES scores may improve while CTBS scores do not. In this case the equating will become outdated. The District plans to have periodic checks and up-date the equating. The exact details of this have not as yet been planned.

Norm-referenced Testing

The equating study will give the District the opportunity to eliminate all norm-referenced testing at the elementary school level; however, there is some question whether this is wise. Many parents like to see how their child is achieving compared with a national norm, and the District needs to see how it is doing compared to these same norms. The District anticipates that it will have a program that will give each child an opportunity to take a norm-referenced test once during elementary school. The details of this program have not been worked out as yet.

APPENDIX

Tables 1-8

Table I

Tests Currently Given Title I Pupils at the Elementary School Level and the Time Required for Each Test

Test Section	GRADE 1	2	3	4	5	6	Time Required (in minutes)
CTBS							
Reading	X	X	X	X	X	X	50
Mathematics	X	X	X	X	X	X	75
SES							
Reading	X	X	X	X	X	X	45
Mathematics	X	X	X	X	X	X	60
Language	X	X	X	X	X	X	45
Written Composition			X			X	45
CAP							
Reading			X			X	
Written Expression			X			X	30 minutes all sections combined.
Mathematics			X			X	
Spelling			X			X	
Entry Level	X						30

Table 2

FREQUENCY DISTRIBUTION FOR SES GRADE 5 READING SCORES

SCORE	NUMBER	CUM NO	PCT	CUM PCT
0	0	0	.00	.00
1	0	0	.00	.00
2	0	0	.00	.00
3	0	0	.00	.00
4	2	2	.24	.24
5	2	4	.24	.47
6	2	6	.24	.71
7	3	9	.36	1.07
8	3	12	.36	1.42
9	4	16	.47	1.90
10	6	22	.71	2.61
11	10	32	1.19	3.80
12	6	38	.71	4.51
13	5	43	.59	5.10
14	9	52	1.07	6.17
15	9	61	1.07	7.24
16	16	77	1.90	9.13
17	8	85	.95	10.08
18	9	94	1.07	11.15
19	7	101	.83	11.98
20	15	116	1.78	13.76
21	11	127	1.30	15.07
22	12	139	1.42	16.49
23	15	154	1.78	18.27
24	11	165	1.30	19.57
25	20	185	2.37	21.95
26	18	203	2.14	24.08
27	22	225	2.61	26.69
28	16	241	1.90	28.59
29	17	258	2.02	30.60
30	21	279	2.49	33.10
31	24	303	2.85	35.94
32	27	330	3.20	39.15
33	32	362	3.80	42.94
34	32	394	3.80	46.74
35	18	412	2.14	48.87
36	41	453	4.86	53.74
37	60	513	7.12	60.85
38	36	549	4.27	65.12
39	59	608	7.00	72.12
40	61	669	7.24	79.36
41	61	730	7.24	86.60
42	51	781	6.05	92.65
43	37	818	4.39	97.03
44	25	843	2.97	100.00

Table 3

FREQUENCY DISTRIBUTION FOR CTBS GRADE 5 READING SCORES

SCORE	NUMBER	CUM NO	PCT	CUM PCT
0	0	0	.00	.00
1	0	0	.00	.00
2	0	0	.00	.00
3	0	0	.00	.00
4	0	0	.00	.00
5	0	0	.00	.00
6	0	0	.00	.00
7	1	1	.12	.12
8	0	1	.00	.12
9	0	1	.00	.12
10	1	2	.12	.24
11	4	6	.48	.71
12	3	9	.36	1.07
13	2	11	.24	1.31
14	5	16	.60	1.90
15	2	18	.24	2.14
16	8	26	.95	3.10
17	9	35	1.07	4.17
18	11	46	1.31	5.48
19	9	55	1.07	6.55
20	16	71	1.90	8.45
21	13	84	1.55	10.00
22	12	96	1.43	11.43
23	18	114	2.14	13.57
24	10	124	1.19	14.76
25	13	137	1.55	16.31
26	21	158	2.50	18.81
27	18	176	2.14	20.95
28	17	193	2.02	22.98
29	15	208	1.79	24.76
30	21	229	2.50	27.26
31	13	242	1.55	28.81
32	15	257	1.79	30.60
33	16	273	1.90	32.50
34	15	288	1.79	34.29
35	15	303	1.79	36.07
36	14	317	1.67	37.74
37	9	326	1.07	38.81
38	15	341	1.79	40.60
39	20	361	2.38	42.98
40	17	378	2.02	45.00
41	12	390	1.43	46.43
42	12	402	1.43	47.86
43	14	416	1.67	49.52
44	9	425	1.07	50.60
45	8	433	.95	51.55
46	15	448	1.79	53.33
47	16	464	1.90	55.24
48	21	485	2.50	57.74
49	15	500	1.79	59.52
50	15	515	1.79	61.31
51	9	524	1.07	62.38

Table 3 (continued)

52	13	537	1.55	63.93
53	10	547	1.19	65.12
54	16	563	1.90	67.02
55	22	585	2.62	69.64
56	16	601	1.90	71.55
57	12	613	1.43	72.98
58	17	630	2.02	75.00
59	10	640	1.19	76.19
60	15	655	1.79	77.98
61	14	669	1.67	79.64
62	9	678	1.07	80.71
63	10	688	1.19	81.90
64	12	700	1.43	83.33
65	15	715	1.79	85.12
66	10	725	1.19	86.31
67	11	736	1.31	87.62
68	14	750	1.67	89.29
69	13	763	1.55	90.83
70	11	774	1.31	92.14
71	10	784	1.19	93.33
72	8	792	.95	94.29
73	8	800	.95	95.24
74	3	803	.36	95.60
75	6	809	.71	96.31
76	4	813	.48	96.79
77	7	820	.83	97.62
78	4	824	.48	98.10
79	6	830	.71	98.81
80	0	830	.00	98.81
81	3	833	.36	99.17
82	3	836	.36	99.52
83	3	839	.36	99.88
84	0	839	.00	99.88
85	1	840	.12	100.00

Table 4

EQUATED SES AND CTBS SCORES FOR GRADE 5 READING

SES	CTBS	CTBS RD
0	.00	0
1	.00	0
2	1.27	1
3	5.16	5
4	9.13	9
5	10.55	11
6	11.33	11
7	12.11	12
8	13.05	13
9	14.19	14
10	15.43	15
11	16.54	17
12	17.39	17
13	18.04	18
14	18.69	19
15	19.49	19
16	20.39	20
17	21.22	21
18	21.88	22
19	22.51	23
20	23.27	23
21	24.18	24
22	25.08	25
23	25.87	26
24	26.61	27
25	27.48	27
26	28.54	29
27	29.68	30
28	30.80	31
29	31.93	32
30	33.16	33
31	34.66	35
32	36.55	37
33	38.55	39
34	40.38	40
35	42.56	43
36	45.24	45
37	48.41	48
38	51.68	52
39	55.20	55
40	59.33	59
41	63.99	64
42	68.66	69
43	73.22	73
44	78.94	79

Table 5

Miscellaneous Necessary Data for the

Equating Study - Grade 5 Reading

Pearson Correlation (SES - CTBS) Corrected for attenuation	.82
K-R 20	
SES	.94
CTBS	.96
Median	
SES	35.73
CTBS	43.94
Mean	
SES	32.53
CTBS	44.61
Variance	
SES	87.57
CTBS	313.24
Standard Deviation	
SES	9.36
CTBS	17.70
Number of Subjects	
SES	843
CTBS	840

Table 6

EQUATING ERROR FOR GRADE 5 READING

SES	CTBS	SE2	SE
0	0	1.60	1.27
1	0	1.52	1.23
2	1	1.43	1.20
3	5	1.35	1.16
4	9	1.26	1.12
5	11	1.19	1.09
6	11	1.11	1.05
7	12	1.04	1.02
8	13	.97	.98
9	14	.90	.95
10	15	.84	.92
11	17	.78	.88
12	17	.72	.85
13	18	.66	.81
14	19	.61	.78
15	19	.56	.75
16	20	.51	.72
17	21	.47	.68
18	22	.43	.65
19	23	.39	.62
20	23	.35	.59
21	24	.32	.56
22	25	.29	.54
23	26	.26	.51
24	27	.23	.48
25	27	.21	.46
26	29	.19	.44
27	30	.18	.42
28	31	.16	.40
29	32	.15	.39
30	33	.14	.38
31	35	.14	.37
32	37	.13	.37
33	39	.13	.37
34	40	.14	.37
35	43	.14	.38
36	45	.15	.39
37	48	.16	.40
38	52	.18	.42
39	55	.19	.44
40	59	.21	.46
41	64	.23	.48
42	69	.26	.51
43	73	.29	.53
44	79	.32	.56

Table 7

MATCHING SES AND CTBS RAW SCORES AND CTBS NATIONAL PERCENTILES FOR GRADE 5 READING

SES	CTBS	NATL %ILE
0	0	1
1	0	1
2	1	1
3	5	1
4	9	1
5	11	1
6	11	1
7	12	1
8	13	1
9	14	1
10	15	1
11	17	2
12	17	2
13	18	3
14	19	3
15	19	3
16	20	4
17	21	5
18	22	6
19	23	6
20	23	6
21	24	7
22	25	9
23	26	10
24	27	11
25	27	11
26	29	14
27	30	15
28	31	17
29	32	18
30	33	20
31	35	23
32	37	26
33	39	29
34	40	31
35	43	36
36	45	39
37	48	44
38	52	50
39	55	54
40	59	60
41	64	67
42	69	77
43	73	85
44	79	96

Table 8

MATCHING SES AND CTBS RAW SCORES AND CTBS NATIONAL PERCENTILES FOR GRADE 1 MATHEMATICS

SES	CTBS	NATL %ILE
0	0	1
1	0	1
2	0	1
3	0	1
4	0	1
5	0	1
6	0	1
7	0	1
8	0	1
9	2	1
10	5	1
11	6	1
12	7	1
13	8	1
14	9	1
15	11	1
16	13	2
17	14	3
18	14	3
19	15	4
20	16	5
21	17	7
22	18	8
23	19	10
24	21	15
25	23	21
26	26	31
27	34	56
28	46	81